DOCUMENT RESUME

ED 264 777 HE 018 951

AUTHOR Verhulst, Steven J.; And Others
TITLE Dimensions of Physician Performance during Residency.
PUB DATE Oct 84
NOTE 17p.; Paper presented to a meeting of the Evaluation Network/Evaluation Research Society (San Francisco, CA, October 10-13, 1984).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Clinical Diagnosis; *Clinical Experience; Competence; *Graduate Medical Students; Higher Education; *Medical Schools; Physician Patient Relationship; *Physicians; *Skill Development; *Student Evaluation

ABSTRACT

Ratings of clinical competence during residency were undertaken for graduates of two medical schools. Near the end of the first year of residency, supervisor ratings were made of 13 variables for 320 graduates in seven classes at Southern Illinois University School of Medicine and one class at the University of Iowa College of Medicine. Principal component analysis with varimax rotations was undertaken for the combined classes, resulting in two factors representing a clinical skills dimension and a professional behavior dimension. Virtually identical factor solutions were obtained for each of the seven classes considered separately and for the class at the University of Iowa, indicating the reliability of the factor structure. Further meaning of the factors was revealed by the pattern of correlations with various comparison measures. These results provide a basis for summarizing supervisor ratings of clinical competence of first-year residents. Additional work regarding the validity of these factors is needed if the instrument is to be used as an outcome measure regarding effectiveness of educational programs or as a predictor of subsequent physician performance. (Author/SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Dimensions of Physician Performance During Residency

Steven J. Verhulst, Jerry A. Colliver, Ph.D.,
Rosalia E.A. Paiva, Ph.D., and Reed G. Williams, Ph.D.

ED264777

COPYRIGHT BOARD OF REGENTS

Address Reprint requests to:

Steven J. Verhulst
Division of Statistics & Measurement
SIU School of Medicine
PO Box 3926
Springfield, Illinois 62708

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Steven J. Verhulst

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

HE018951

Mr. Verhulst is an Instructor and Dr. Colliver is Visiting Professor and Director of the Division of Statistics & Measurement. Dr. Paiva (now deceased) was Associate Professor and Dr. Williams is Professor and Chairman of the Department of Medicial Education; all at Southern Illinois University School of Medicine.

Dimensions of Physician Performance During Residency

Steven J. Verhulst, Jerry A. Colliver, Ph.D.,
Rosalia E.A. Paiva, Ph.D., and Reed G. Williams, Ph.D.

Abstract

Supervisor ratings of 13 variables depicting clinical competence were obtained for 320 graduates of seven classes of Southern Illinois University School of Medicine and one class from the University of Iowa College of Medicine near the end of the graduates' first year of residency. Principal components analysis with varimax rotations of the data for the classes combined resulted in two factors representing a Clinical Skills dimension and a Professional Behavior dimension. Virtually identical factor solutions were obtained for each of the seven classes considered separately and for the class at the University of Iowa, thus indicating the reliability of the factor structure. Further meaning of the factors was revealed by the pattern of correlations with various comparison measures. These results provide a basis for summarizing supervisor ratings of clinical competence of first year residents. Additional work regarding the validity of these factors is needed if the instrument is to be used as an outcome measure regarding effectiveness of educational programs or as a predictor of subsequent physician performance.

Dimensions of Physician Performance During Residency

Steven J. Verhulst, Jerry A. Colliver, Ph.D.,
Rosalia E.A. Paiva, Ph.D., and Reed G. Williams, Ph.D.
Southern Illinois University School of Medicine

The search for adequate indicators of medical performance has occupied the field of medical education for years. Such indicators are needed as outcome measures against which to validate examinations such as the MCAT and NBME exams and to validate medical education programs. They are also needed as predictors of future physician performance for selection purposes of residency directors and future employers.

Various instruments have been used for these purposes. Most have been some form of rating scale. Factor analytic studies have been performed to determine the nature of such evaluations of medical students' and residents' during their medical training. The most common finding of these studies is that the evaluations are reflecting two underlying dimensions which represent a knowledge or cognitive factor and a non-cognitive factor. Schumacher, for example, identified two orthogonal factors underlying eleven measures including medical school grades, National Board examination scores and senior year peer ratings of 306 students participating in the AAMC Longitudinal Study; a "General medical knowledge" factor defined by grades and scores, and a "skill in patient relationships" factor defined by the peer ratings (1).

Gough and his colleagues found three orthogonal factors -- "medical competence," "medical identity" and "medical effectiveness" -- in their factor analysis of faculty ratings of 81 third-year clerks on ten measures, but only two factors -- "general competence" and "effectiveness in practice" -- when the students were rated on the same variables during their senior year (2). Geertsma and Chapman identified two common and six specific obliquely rotated factors accounting for the intercorrelations between college grade point average, Medical College Admission Test scores and 14 evaluative ratings of 180 medical students: the two common factors were interpreted as a general, mostly cognitive factor and a non-cognitive dimension (3). More recently, Dielman, Hull and Davis (4) also extracted two correlated factors termed "problem solving" and "interpersonal skills" from 15 ratings of performance of junior and senior medical students, and replicated the finding in twelve consecutive instructional periods on samples varying from 107 to 186 subjects.

Evaluation of medical school graduates by their residency supervisors can provide important information for judging the adequacy of medical school preparation for subsequent specialty training. For this reason, the Southern Illinois University School of Medicine initiated a residency follow-up study during the School's early years. In preparation for the study, specific components of residency performance were identified, which teaching faculty considered of major importance for evaluating the achievement of the goals of the medical school curriculum. From these, an evaluation form was developed to gather supervisor ratings of performance in 13 specific areas. Thus, since 1976, residency supervisors have been asked to rate SIU School of Medicine

graduates in their residency programs near the end of the first year of residency. Upon completion of the data collection for each class, the information is compiled and distributed to School of Medicine faculty.

The purpose of the present study was to identify the factor structure of supervisor ratings of house officers' performance and to assess the reliability of that structure through seven classes of SIU School of Medicine graduates. It was hoped that the results of the study would suggest a meaningful summary of supervisor ratings for presentation in an annual report to faculty.

METHODS

Questionnaire. A 13-item questionnaire was constructed which reflected specific components of residency performance considered by faculty to be of central importance for evaluating the achievement of the goals of the medical school curriculum (See Table 1 for summary of 13 items.) Development of these items was based on studies of clinical competence; especially the critical incident study of first year residents completed by the American Institutes for Research and commissioned by the NBME (5). Four additional items were included to assess overall, comprehensive performance. These items asked about overall competence in medicine in general, overall competence in the specialty, competence in comparison with other residents, and desirability as personal physician. Supervisors were asked to rate residents on a 10-point scale ranging from "1" (unsatisfactory) to "10" (outstanding). Studies

regarding features of the instrument including inter rater reliability has been reported previously (6).

Insert Table 1 about here

Subjects. Questionnaires were mailed to the residency supervisors of the 440 SIU graduates of the seven classes from 1976 to 1982 near the end of the first year of residency. Supervisor ratings on the questionnaire for graduates of the 1977 class of the University of Iowa College of Medicine were available for cross validating the SIU findings.

Statistical Analyses. A principal components analysis with varimax rotations (7, 8) was performed on the ratings on the 13 specific items for the seven classes combined and for each of the seven classes separately. The reliability of the factors was assessed by determining the agreement among the factor structures for the seven classes. Agreement was ascertained by visual inspection and by assessment with coefficients of congruence (7). Further meaning of the factors was assessed by correlating factor scores with the overall, comprehensive ratings of performance during residency and with scores on medical licensure examinations (NBME, FLEX, etc.).

RESULTS

Questionnaires with supervisor ratings were returned for 320 of the 440 SIU graduates (73%) of the seven classes from 1976 to 1982. The number and percentage of responding supervisors for each of the seven classes are presented in Table 1, which shows that return rates ranged from 58% for the class of 1978 to 90% for the class of 1982. Supervisor ratings were available for 44 out of 169 graduates (26%) of the 1977 class of the University of Iowa College of Medicine.

The principal components analysis with varimax rotations performed on the data for all 320 graduates of the seven classes combined resulted in two factors which accounted for 84% of the total aggregate variance (see Table 1). The eight variables loading most heavily on the first factor were history-taking, physical exam, problem identification, diagnostic approach, management, knowledge, independent learning habits, and technical skills, which connotatively defined a Clinical Skills factor. The five variables loading most heavily on the second factor were relationship with patients, relationship with medical staff, reliability, self-awareness, and emotional stability, which defined a Professional Behavior factor. The associations among items reflected by the factor analysis are consistent with common conceptions of clinical competence. The two factors accounted for 49% and 34%, respectively, of the total aggregate variance.

A separate principal components analysis with varimax rotations was performed on the data for each of the seven classes. The factor solution for each of the seven classes was virtually identical to that obtained for the seven classes combined (see Table 1). The analysis for each class resulted in two factors, and the pattern of factor loadings of the 13 variables on the two factors was generally identical to that for the seven classes combined. The total aggregate variance accounted for by the two factors together ranged from 79% for the classes of 1976 and 1978 to 89% for the class of 1980. For the Clinical Skills factor, the variance accounted for ranged from 44% to 55% and for the Profesional Behavior factor, from 27% to 38%.

Coefficients of congruence were used as an objective supplement to visual inspection to assess the factorial similarity of the results obtained for the seven classes separately. (See Table 2.) The coefficients were quite high ranging from 0.91 to 0.99 for the Clinical Skills factor and from 0.87 to 0.99 for the Professional Behavior factor.

Insert Table 2 about here

Further support for the reliability of the factor structure was provided by results obtained from an analysis of residency supervisor ratings of 1977 graduates of the University of Iowa College of Medicine (See Table 1). The results were virtually identical to those obtained for the seven SIU classes.

Ratings of overall competence and ratings of comparison with other residents correlated slightly higher with scores on the Clinical Skills factor than with scores on the Professional Behavior factor, indicating that supervisors see Clinical Skills as a more important component of overall competence than Professional Behavior. (See Table 3.) On the other hand, there was virtually no difference in the correlations of ratings of desirability as a personal physician and the two factors, suggesting that supervisors see Professional Behavior and Clinical Skills as equally important in choice of a personal physician.

Insert Table 3 about here

Scores on national licensure examinations (NBME, FLEX, etc.) designed to measure medical knowledge had small, positive correlations with the Clinical Skills factor, indicating the contribution of medical knowledge to the factor. (See Table 3.) These exam results revealed small, negative correlations with the Professional Behavior factor. As anticipated, the correspondence between supervisor ratings and NBME Parts II and III scores are more substantial than that with NBME Part I.

DISCUSSION

Principal components analyses with varimax rotations of supervisor ratings on thirteen variables resulted in two factors, a Clinical Skills factor and a Professional Behavior factor. The factor structures for the seven classes combined and for each class separately were virtually identical. Similar results were obtained for the 1977 class of the University of Iowa College of Medicine. The factorial simularity was confirmed by high coefficients of congruence and the meaning of the factors was reinforced by the pattern of correlations of the factor scores with NBME and FLEX scores.

The importance of these results is their implication that residency performance can be meaningfully summarized in terms of two dimensions, Clinical Skills and Professional Behavior, which are not simply duplicating the traditional cognitive medical knowledge dimension(s) reflected in certifying examination scores. Additional studies to further clarify our understanding of the meaning of these scores are planned. Results so far suggest that the instrument will provide a simple, direct method of evaluating clinical competence, as defined by the judgement of experts. For research purposes, the instrument will provide outcome measures to assess the effectiveness of educational interventions and predictor measures to anticipate subsequent residency and eventual practice performance.

3702m

REFERENCES

1. Schumacher, C.F. A factor-analytic study of various criteria of medical student accomplishment. Journal of Medical Education, 39:192-196, 1964.

2. Gough, H.G., Hall, W.B., and Harris, R.E. Evaluation of performance in medical training. Journal of Medical Education, 39:679-692, 1964.

3. Geertsma, R.H., and Chapman, J.E. The evaluation of medical students. Journal of Medical Education, 42:938-948, 1967.

4. Dielman, T.E., Hull, A.L., and Davis, W.K. Psychometric properties of clinical performance ratings. Evaluation and the Health Professions, 3:103-117, 1980.

5. American Institutes for Research. The definition of clinical competence in Medicine: Performance dimensions and rationale for clinical skill area. Palo Alto, CA: American Institute for Research in the Behavioral Science, 1976.

6. Paiva, R.E.A. Reliability in follow-up evaluations of graduates as residents. Presented at the annual meeting of the Association of American Medical Colleges, Washington, D.C., November, 1979.

7. Harman, H.H. Modern Factor Analysis, Second Edition. Chicago, The University of Chicago Press, 1967.

8. SAS Institute Inc. SAS User's Guide: Statistics, 1982 Edition. Cary, NC: SAS Institute, Inc., 1982.

Clinical Skills	1976	1977	1978	1979	1980	1981	1982	Combined	Iowa
1. History	.83	.79	.90	.65	.76	.73	.86	.81	.82
2. Physical Exam	.84	.86	.88	.70	.81	.84	.83	.84	.81
3. Problem Identification	.93	.85	.91	.70	.87	.90	.86	.87	.86
4. Diagnostic Approach	.81	.81	.93	.68	.80	.75	.81	.81	.72
5. Management	.73	.67	.92	.38	.84	.85	.71	.77	.76
6. Knowledge	.89	.90	.81	.89	.86	.90	.88	.89	.83
7. Independent Learning Habits	.73	.85	.64	.90	.79	.86	.70	.78	.67
8. Relationship/Patients	.14	.34	.72	.09	.36	.37	.35	.39	.31
9. Relationship/Staff	.01	.28	.12	.40	.47	.22	.31	.28	.20
10. Reliability	.33	.43	.24	.66	.65	.55	.51	.52	.32
11. Self-awareness	.60	.62	.56	.73	.79	.69	.58	.65	.49
12. Emotional Stability	.38	.38	.31	.46	.60	.40	.40	.42	.52
13. Technical Skills	.86	.67	.69	.75	.82	.79	.68	.75	.80
% Variance	47%	47%	51%	44%	55%	51%	48%	49%	44%
Professional Behavior									
1. History	.34	.46	.17	.62	.51	.56	.42	.43	.37
2. Physical Exam	.34	.34	.33	.60	.44	.41	.42	.40	.28
3. Problem Identification	.17	.42	.27	.61	.44	.30	.42	.38	.32
4. Diagnostic Approach	.44	.48	.22	.52	.52	.50	.49	.46	.50
5. Management	.55	.66	.23	.72	.51	.41	.60	.53	.41
6. Knowledge	.08	.26	.44	.28	.45	.32	.36	.31	.38
7. Independent Learning Habits	.26	.34	.56	.18	.33	.24	.47	.33	.41
8. Relationship/Patients	.89	.81	.42	.92	.90	.84	.72	.79	.82
9. Relationship/Staff	.93	.92	.86	.77	.83	.94	.91	.91	.94
10. Reliability	.79	.82	.78	.62	.68	.73	.77	.73	.79
11. Self-awareness	.69	.68	.64	.54	.49	.65	.72	.65	.72
12. Emotional Stability	.74	.95	.79	.77	.69	.86	.80	.81	.69
13. Technical Skills	.05	.50	.39	.30	.49	.49	.61	.46	.15
% Variance	32%	38%	27%	37%	34%	36%	38%	34%	33%
Total % Variance	79%	85%	79%	81%	89%	87%	86%	84%	76%
Supervisors Responding:	36	39	38	43	44	55	65	320	44
Class Size:	43	47	57	74	75	72	72	440	169
%:	84%	83%	67%	58%	59%	76%	90%	73%	26%

Table 1. Factor loadings of 13 variables on Clinical Skills Factor and Professional Behavior Factor. To emphasize the similar factor structure for the seven classes separately and combined and for the Iowa class, the larger loading of a variable on the two factors is underlined.

	Clinical Skills Factor							Professional Behavior Factor						
Class	1977	1978	1979	1980	1981	1982	IOWA	1977	1978	1979	1980	1981	1982	IOWA
1976	0.98	0.96	0.96	0.97	0.98	0.97	0.99	0.97	0.91	0.94	0.94	0.96	0.94	0.98
1977		0.97	0.98	0.99	0.99	0.99	0.99		0.94	0.97	0.98	0.99	0.99	0.98
1978			0.91	0.95	0.97	0.98	0.98			0.87	0.92	0.94	0.96	0.95
1979				0.99	0.98	0.96	0.96				0.98	0.96	0.96	0.94
1980					0.99	0.99	0.98					0.99	0.98	0.97
1981						0.99	0.99						0.98	0.98
1982							0.99							0.97

Table 2. Coefficients of Congruence

Comparison Measures	n	Clinical Skills Factor	Professional Behavior Factor
Overall Competence/Medicine	220	0.84***	0.48***
Overall Competence/Specialty	219	0.78***	0.56***
Overall Competence/Average	221	0.82***	0.53***
Comparison with other house staff	221	0.77***	0.51***
Desirability as Personal Physician	219	0.68***	0.63***
FLEX Day 1	134	0.20*	-0.18*
FLEX Day 2	134	0.17	-0.20*
FLEX Day 3	134	0.23**	-0.14
FLEX Weighted Average	134	0.22**	-0.18*
NBME Part I	101	0.18	-0.12
NBME Part II	91	0.23*	-0.24*
NBME Part III	61	0.33*	-0.15

* $p < .05$
** $p < .01$
*** $p < .001$

Table 3. Correlations between Standardized Factor Scores and Comparison Measures.